



Searching within The ACM Digital Library with Advanced Search: (crawler) ([start a new search](#))

Found 1,040 of 239,964

REFINE YOUR SEARCH

▼ Refine by Keywords

[Discovered Terms](#)

▼ Refine by People

[Names](#)
[Institutions](#)
[Authors](#)
[Editors](#)

▼ Refine by Publications

[Publication Year](#)
[Publication Names](#)
[ACM Publications](#)
[All Publications](#)
[Content Formats](#)
[Publishers](#)

▼ Refine by Conferences

[Sponsors](#)
[Events](#)
[Proceeding Series](#)

ADVANCED SEARCH

[Advanced Search](#)

FEEDBACK

[Please provide us with feedback](#)

Found 1,040 of 239,964

 [Search Results](#) [Related Journals](#) [Related Magazines](#)
 [Related SIGs](#) [Related Conferences](#)

Results 1 - 20 of 1,040

Sort by

in

 [Save results to a](#)
[Binder](#)

Result page: 1 2 3 4 5 6 7 8 9 10 next >>

1 [Distributed monitoring of peer to peer systems](#)
[Serge Abiteboul, Bogdan Marinou](#)
November WIDM '07: Proceedings of the 9th annual ACM international workshop on Web information and data management
2007

Publisher: ACM

Full text available: [PDF](#) (509.10 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index](#), [terms](#)**Bibliometrics:** Downloads (6 Weeks): 8, Downloads (12 Months): 177, Citation Count: 0

In this paper, we are concerned with the distributed monitoring of P2P systems. We introduce the P2P Monitor system and a new declarative language, namely P2PML, for specifying monitoring tasks. A P2PML subscription is compiled into a distributed algebraic ...

Keywords: active documents, databases, distributed data management, peer to peer systems, stream processing, web services

2 [What's going on?: learning communication rules in edge networks](#)
[Srikanth Kandula, Ranveer Chandra, Dina Katabi](#)

August SIGCOMM '08: Proceedings of the ACM SIGCOMM 2008 conference on 2008 Data communication

Publisher: ACM

Full text available: [PDF](#) (4.84 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [index](#), [terms](#)**Bibliometrics:** Downloads (6 Weeks): 46, Downloads (12 Months): 135, Citation Count: 0

Existing traffic analysis tools focus on traffic volume. They identify the heavy-hitters - flows that exchange high volumes of data, yet fail to identify the structure implicit in network traffic - do certain flows happen before, after or along with ...

Keywords: communication rules, correlation, expose, rule mining

Also published in:

October 2008 SIGCOMM Computer Communication Review Volume 38 Issue 4

3 Automatically constructing collections of online database directories

 Luciano Barbosa, Juliana Freire

November CIKM '06: Proceedings of the 15th ACM international conference on 2006 Information and knowledge management

Publisher: ACM

Full text available:  [Pdf](#) (281.08 KB)

Additional Information: [full citation](#), [references](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 6, Downloads (12 Months): 54, Citation Count: 0

Keywords: focused web crawling, modular classifiers, online databases

4 Creating a Web community chart for navigating related communities

 Masashi Toyoda, Masaru Kitsuregawa

September HYPERTEXT '01: Proceedings of the 12th ACM conference on 2001 Hypertext and Hypermedia

Publisher: ACM

Full text available:  [Pdf](#) (276.51 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [cited by](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 9, Downloads (12 Months): 66, Citation Count: 11

Recent research on link analysis has shown the existence of numerous web communities on the Web. A web community is a collection of web pages created by individuals or any kind of associations that have a common interest on a specific topic. In this ...

Keywords: Web community, World Wide Web, link analysis, related web communities

5 Exploring traversal strategy for web forum crawling

 Yida Wang, Jiang-Ming Yang, Wei Lai, Rui Cai, Lei Zhang, Wei-Ying Ma

July SIGIR '08: Proceedings of the 31st annual international ACM SIGIR 2008 conference on Research and development in information retrieval

Publisher: ACM

Full text available:  [Pdf](#) (522.27 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 45, Downloads (12 Months): 200, Citation Count: 0

In this paper, we study the problem of Web forum crawling. Web forum has now become an important data source of many Web applications; while forum crawling is still a challenging task due to complex in-site link structures and login controls of most ...

Keywords: forum crawler, sitemap, traversal strategy

6 People-centric urban sensing

 [Andrew T. Campbell, Shane B. Eisenman, Nicholas D. Lane, Emiliano Miluzzo, Ronald A. Peterson](#)
 August 2006 WICON '06: Proceedings of the 2nd annual international workshop on Wireless internet
 Publisher: ACM

Full text available:  [Pdf](#) (249.73 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 44, Downloads (12 Months): 268, Citation Count: 8

The vast majority of advances in sensor network research over the last five years have focused on the development of a series of small-scale (100s of nodes) testbeds and specialized applications (e.g., environmental monitoring, etc.) that are built on ...

7 Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval

 [Tat-Seng Chua, Mun-Kew Leong, Syung Hyon Myaeng, Douglas W. Oard, Fabrizio Sebastiani](#)
 July 2008 SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval
 Publisher: ACM

Additional Information: [full citation](#), [abstract](#)

Bibliometrics: Downloads (6 Weeks): n/a, Downloads (12 Months): n/a, Citation Count: 0

Welcome to the 31st year of SIGIR, the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. The growth in SIGIR in recent years has been remarkable. SIGIR 2005 received a record 368 full paper submissions, SIGIR ...

8 Robust mesh watermarking

[Emil Praun, Hugues Hoppe, Adam Finkelstein](#)
 July 1999 SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques
 Publisher: ACM Press/Addison-Wesley Publishing Co.

Full text available:  [Pdf](#) (2.08 MB) Additional Information: [full citation](#), [references](#), [cited by](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 25, Downloads (12 Months): 126, Citation Count: 17

Keywords: copyright protection, steganography

9 Extended static checking for Java

 Cormac Flanagan, K. Rustan M. Leino, Mark Lillibridge, Greg Nelson, James B. Saxe, Baymie Stata

June PLDI '02: Proceedings of the ACM SIGPLAN 2002 Conference on
2002 Programming language design and implementation

Publisher: ACM

Full text available:  Pdf (257.73 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [cited by](#),
[index terms](#)

Bibliometrics: Downloads (6 Weeks): 37, Downloads (12 Months): 257, Citation Count: 138

Software development and maintenance are costly endeavors. The cost can be reduced if more software defects are detected earlier in the development cycle. This paper introduces the Extended Static Checker for Java (ESC/Java), an experimental compile-time ...

Keywords: compile-time program checking

Also published in:

May 2002 **SIGPLAN Notices** Volume 37 Issue 5

10 On the responsiveness of DNS-based network control

 Jeffrey Pang, Aditya Akella, Anees Shaikh, Balachander Krishnamurthy, Srinivasan Seshan

October IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet
2004 measurement

Publisher: ACM

Full text available:  Pdf (255.96 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [cited by](#),
[index terms](#)

Bibliometrics: Downloads (6 Weeks): 11, Downloads (12 Months): 53, Citation Count: 4

For the last few years, large Web content providers interested in improving their scalability and availability have increasingly turned to three techniques: mirroring, content distribution, and ISP multihoming. The Domain Name System (DNS) has gained ...

Keywords: DNS, network control, time-to-live

11 Proceeding of the 10th ACM workshop on Web information and data

 management

Chee-Yong Chan, Neoklis Polyzotis

October WIDM '08: Proceeding of the 10th ACM workshop on Web information
2008 and data management

Publisher: ACM

Additional Information: [full citation](#), [abstract](#)

Bibliometrics: Downloads (6 Weeks): n/a, Downloads (12 Months): n/a, Citation Count: 0

The ACM CIKM08 Workshop on Web Information and Data Management (WIDM 2008) is the tenth in a series of workshops on Web Information and Data Management held in conjunction with the International Conference on Information and Knowledge Management ...

12 A probabilistic model for retrospective news event detection

Zhiwei Li, Bin Wang, Mingjing Li, Wei Ying Ma

August 2005 SIGIR '05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval

Publisher: ACM

Full text available: Pdf (435.47 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [cited by](#), [index terms](#)**Bibliometrics:** Downloads (6 Weeks): 14, Downloads (12 Months): 152, Citation Count: 9

Retrospective news event detection (RED) is defined as the discovery of previously unidentified events in historical news corpus. Although both the contents and time information of news articles are helpful to RED, most researches focus on the utilization ...

Keywords: clustering, expectation maximization, maximum likelihood, retrospective news event detection

13 Combined searching of web and oai digital library resources

Aaron Krowne, Martin Halbert

June 2004 JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries

Publisher: ACM

Full text available: Pdf (77.07 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [cited by](#), [index terms](#)**Bibliometrics:** Downloads (6 Weeks): 5, Downloads (12 Months): 21, Citation Count: 1

In this paper, we describe an experiment in combined searching of web pages and digital library resources, exposed via an Open Archives metadata provider and web gateway service. We utilize only free/open source software components for our investigation, ...

Keywords: DP9, OAI, crawlers, digital libraries, open archives, search engines

14 Finding near-duplicate web pages: a large-scale evaluation of algorithms

Monika Henzinger

August 2006 SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval

Publisher: ACM

Full text available: Pdf (312.35 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)**Bibliometrics:** Downloads (6 Weeks): 58, Downloads (12 Months): 403, Citation Count: 10

Broder et al.'s [3] shingling algorithm and Charikar's [4] random projection based approach are considered "state-of-the-art" algorithms for finding near-duplicate web pages. Both algorithms were either developed at or used by popular web search engines. ...

Keywords: content duplication, near-duplicate documents, web pages

15 Querying for meta knowledge

 [Bernhard Schueler](#), [Sergei Sizov](#), [Steffen Staab](#), [Duc Thanh Tran](#)

April 2008 WWW '08: Proceeding of the 17th international conference on World Wide Web
Publisher: ACM

Full text available:  [Pdf](#) (217.47 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 40, Downloads (12 Months): 186, Citation Count: 0

The Semantic Web is based on accessing and reusing RDF data from many different sources, which one may assign different levels of authority and credibility. Existing Semantic Web query languages, like SPARQL, have targeted the retrieval, combination ...

Keywords: RDF, SPARQL, semantic web

16 How dynamic are IP addresses?

 [Yinglian Xie](#), [Fang Yu](#), [Kannan Achan](#), [Eliot Gilum](#), [Moises Goldszmidt](#), [Ted Wobber](#)

August 2007 SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications

Publisher: ACM

Full text available:  [Pdf](#) (553.84 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 32, Downloads (12 Months): 326, Citation Count: 2

This paper introduces a novel algorithm, *UDmap*, to identify dynamically assigned IP addresses and analyze their dynamics pattern. UDmap is fully automatic, and relies only on application-level server logs. We applied UDmap to a month-long Hotmail ...

Keywords: DHCP, IP volatility, dynamic IP addresses, entropy, spam detection

Also published in:

October 2007 **SIGCOMM Computer Communication Review** Volume 37 Issue 4

17 Current trends in web data analysis

 [Arun Sen](#), [Peter A. Dacin](#), [Christos Pattichis](#)

November 2006 **Communications of the ACM** , Volume 49 Issue 11

Publisher: ACM

Full text available:  [Html](#) (31.44 KB),  [Pdf](#) (186.80 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Bibliometrics: Downloads (6 Weeks): 62, Downloads (12 Months): 540, Citation Count: 3

Considering potential reasons for the underutilization of clickstream data and suggesting ways to enhance its use.

18 A Method for Focused Crawling Using Combination of Link Structure and Content Similarity**Mohsen Jamali, Hassan Sayyadi, Babak Bagheri Hariri, Hassan Aboihassani**

December 2006 WI '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence

Publisher: IEEE Computer SocietyFull text available:  [Pdf](#) (3.19 MB)Additional Information: [full citation](#), [abstract](#), [index terms](#)**Bibliometrics:** Downloads (6 Weeks): 4, Downloads (12 Months): 54, Citation Count: 0

The rapid growth of the world-wide web poses unprecedented scaling challenges for general-purpose crawlers and search engines. A focused crawler aims at selectively seek out pages that are relevant to a pre-defined set of topics. Besides specifying topics ...

19 Characterization of the Thai hostgraph**Kulwadee Sambooniyati, Masashi Toyoda, Shinji Suzuki, Masaru Kitsuregawa**

January 2008 ICIIMC '08: Proceedings of the 2nd international conference on Ubiquitous information management and communication

Publisher: ACMFull text available:  [Pdf](#) (283.41 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)**Bibliometrics:** Downloads (6 Weeks): 8, Downloads (12 Months): 38, Citation Count: 0

The Web of a country or the national Web is a set of web pages related to a specific country. Understanding in the graph structure of the national Web provides invaluable insights for the development of algorithms and localized search services targeting ...

20 Keywords: web characterization, web measurement
Simulating non-scanning worms on peer-to-peer networks**Guanling Chen, Robert S. Gray**

May 2006 InfoScale '06: Proceedings of the 1st international conference on Scalable information systems

Publisher: ACMFull text available:  [Pdf](#) (226.74 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)**Bibliometrics:** Downloads (6 Weeks): 16, Downloads (12 Months): 138, Citation Count: 0

Millions of Internet users are using large-scale peer-to-peer (P2P) networks to share content files today. Many other mission-critical applications, such as Internet telephony and Domain Name System (DNS), have also found P2P networks appealing due to ...

Result page: 1 [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

>>

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2008 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)